

CLAIMS

What is claimed is:

1. A method of discovering the elements in a compound with a fitness function, comprising:
 - receiving a set of monoisotopic mass look-up-tables (LUTs) that associates LUT addresses with mass values for a set of elements;
 - identifying mass values in parallel by cross-referencing two or more LUT addresses associated with an electronic chromosome and addresses in the monoisotopic mass LUTs;
 - evaluating different permutations of the identified mass values from the set of monoisotopic mass LUTs;
 - accessing values in two or more mass spectroscopy data sets according to the permutations of mass values identified in the monoisotopic mass LUTs; and
 - determining the combination of elements in the compound according to a correlation between the permutations of mass values and the mass values associated with the mass spectroscopy data set.
2. The method of claim 1 wherein evaluating different permutations further comprises:
 - combining one or more mass values from the set of monoisotopic mass LUTs together; and
 - adding one or more constant values to the combination of one or more mass values.
3. The method of claim 1 wherein the compound is a protein and the elements include one or more amino acids to be discovered during protein sequencing analysis.
4. The method of claim 2 wherein the one or more amino acids are selected from a set of

amino acids including: Gly, Ala, Ser, Pro, Val, Thr, Cys, Ile, Leu, Asn, Asp, Gln, Glu, Met, His, Phe, Arg, Tyr, Trp and Lys.

5. The method of claim 1 wherein discovering elements of a compound relates to aspects of combinatorial chemistry and genomics.

6. The method of claim 1 wherein discovering elements of a compound relates to aspects of medicinal chemistry and high-throughput screening used in conjunction discovering medicinal chemical compounds.

7. The method of claim 1 wherein evaluating different permutations may include one or more of the following: adding or subtracting constants values, summing one or more different mass values.

8. The method of claim 1 wherein the electronic chromosome is a storage area with a series of addresses corresponding to mass values and potentially the elements in the compound.

9. A method of discovering elements of a compound using genetic algorithm computation, comprising:

receiving a population of electronic chromosomes each representing a set of elements and having a fitness value determined using a fitness function indicating the probability that at least one of the electronic chromosome provides the elements in the compound;

selecting at least two electronic chromosomes randomly from the population of electronic chromosomes as parent electronic chromosomes;

creating a child electronic chromosome through a crossover of the at least two parent electronic chromosomes using one or more randomly selected cut-points on the parent

electronic chromosomes; and

evaluating the fitness value of the resulting child electronic chromosome using the fitness function to determine if elements in the child electronic chromosome correspond to the compound.

10. The method of claim 9, further comprising:

determining if the child electronic chromosome enters the population of electronic chromosomes when the fitness value indicates that the elements associated with the child electronic chromosome more likely represents the elements in the compound than the an electronic chromosome already in the population of electronic chromosomes.

11. The method of claim 9 further comprising:

mutating a child electronic chromosome through a probabilistic alteration of the bits representing the child electronic chromosome and a potential combination of elements in the compound.

12. The method of claim 9 wherein the compound is a protein and the elements include one or more amino acids to be discovered during protein sequencing analysis.

13. The method of claim 12 wherein the one or more amino acids are selected from a set of amino acids including: Gly, Ala, Ser, Pro, Val, Thr, Cys, Ile, Leu, Asn, Asp, Gln, Glu, Met, His, Phe, Arg, Tyr, Trp and Lys.

14. The method of claim 9 wherein discovering elements of a compound relates to aspects of combinatorial chemistry and genomics.

15. The method of claim 9 wherein discovering elements of a compound relates to aspects of medicinal chemistry and high-throughput screening used in conjunction

discovering medicinal chemical compounds.

16. The method of claim 9 wherein the electronic chromosome is a storage area with a series of addresses corresponding to mass values and potential elements in the compound.

17. The method of claim 9 wherein the fitness function, comprises

receiving a set of monoisotopic mass look-up-tables (LUTs) that associates LUT addresses with mass values for a set of elements;

identifying mass values in parallel by cross-referencing two or more LUT addresses associated with an electronic chromosome and addresses in the monoisotopic mass LUTs;

evaluating different permutations of the identified mass values from the set of monoisotopic mass LUTs;

accessing values in two or more mass spectroscopy data sets according to the permutations of mass values identified in the LUT; and

determining the combination of elements in the compound according to a correlation between the permutations of mass values and the mass values associated with the mass spectroscopy data set.

18. A computer program product for discovering the elements in a compound with a fitness function, tangibly stored on a computer-readable medium, comprising instructions operable to cause a programmable processor to:

receive a set of monoisotopic mass look-up-tables (LUTs) that associates LUT addresses with mass values for a set of elements;

identify mass values in parallel by cross-referencing two or more LUT addresses associated with an electronic chromosome and addresses in the monoisotopic mass LUTs;

evaluate different permutations of the identified mass values from the set of
monoisotopic mass LUTs;

access values in two or more mass spectroscopy data sets according to the
permutations of mass values identified in the monoisotopic mass LUTs; and

determine the combination of elements in the compound according to a correlation
between the permutations of mass values and the mass values associated with the mass
spectroscopy data set.

19. The computer program product of claim 18 wherein the compound is a protein and
the elements include one or more amino acids to be discovered during protein sequencing
analysis.

20. The computer program product of claim 18 wherein evaluating different permutations
may include one or more of the following: adding or subtracting constants values, summing
one or more different mass values.

21. A computer program product for discovering elements of a compound using genetic
algorithm computation, tangibly stored on a computer-readable medium, comprising
instructions operable to cause a programmable processor to:

receive a population of electronic chromosomes each representing a set of elements
and having a fitness value determined using a fitness function indicating the probability that
at least one of the electronic chromosome provides the elements in the compound;

select at least two electronic chromosomes randomly from the population of
electronic chromosomes as parent electronic chromosomes;

create a child electronic chromosome through a crossover of the at least two parent

electronic chromosomes using one or more randomly selected cut-points on the parent electronic chromosomes; and

evaluate the fitness value of the resulting child electronic chromosome using the fitness function to determine if elements in the child electronic chromosome correspond to the compound.

22. The computer program product of claim 21, further comprising instructions to:

determine if the child electronic chromosome enters the population of electronic chromosomes when the fitness value indicates that the elements associated with the child electronic chromosome more likely represents the elements in the compound than the an electronic chromosome already in the population of electronic chromosomes.

23. The computer program product of claim 21, further comprising instructions to:

mutate a child electronic chromosome through a probabilistic alteration of the bits representing the child electronic chromosome and a potential combination of elements in the compound.

24. The computer program product of claim 21, wherein the compound is a protein and the elements include one or more amino acids to be discovered during protein sequencing analysis.

25. An apparatus for discovering the elements in a compound with a fitness function, comprising:

a set of monoisotopic mass look-up-tables (LUTs) that associates LUT addresses with mass values for a set of elements;

a set of registers operatively connected to the set of monoisotopic mass LUTs that identifies mass values in parallel by cross-referencing two or more LUT addresses associated with an electronic chromosome and addresses in the monoisotopic mass LUTs;

arithmetic logic that evaluates different permutations of the identified mass values from the set of monoisotopic mass LUTs;

storage driven by arithmetic logic to access values in two or more mass spectroscopy data sets according to the permutations of mass values identified in the monoisotopic mass LUTs; and

arithmetic logic that determines the combination of elements in the compound according to a correlation between the permutations of mass values and the mass values associated with the mass spectroscopy data set.

26. An apparatus for discovering elements of a compound using genetic algorithm computation, comprising:

a population of electronic chromosomes in storage each representing a set of elements and having a fitness value determined using a fitness function that indicates the probability that at least one of the electronic chromosome provides the elements in the compound;

selection logic operatively coupled to the population of electronic chromosomes in storage that selects at least two electronic chromosomes randomly from the population of electronic chromosomes as parent electronic chromosomes;

a child electronic chromosome register that receives a crossover of the two parent electronic chromosomes using one or more randomly selected cut-points on the parent electronic chromosomes; and

compound fitness function logic that evaluates the fitness value of the resulting child electronic chromosome using the fitness function to determine if elements in the child electronic chromosome correspond to the compound.

27. An apparatus for discovering the elements in a compound with a fitness function, comprising:

means for receiving a set of monoisotopic mass look-up-tables (LUTs) that associates LUT addresses with mass values for a set of elements;

means for identifying mass values in parallel by cross-referencing two or more LUT addresses associated with an electronic chromosome and addresses in the monoisotopic mass LUTs;

means for evaluating different permutations of the identified mass values from the set of monoisotopic mass LUTs;

means for accessing values in two or more mass spectroscopy data sets according to the permutations of mass values identified in the monoisotopic mass LUTs; and

means for determining the combination of elements in the compound according to a correlation between the permutations of mass values and the mass values associated with the mass spectroscopy data set.

28. An apparatus for discovering elements of a compound using genetic algorithm computation, comprising:

means for receiving a population of electronic chromosomes each representing a set of elements and having a fitness value determined using a fitness function indicating the probability that at least one of the electronic chromosome provides the elements in the

compound;

means for selecting at least two electronic chromosomes randomly from the population of electronic chromosomes as parent electronic chromosomes;

means for creating a child electronic chromosome through a crossover of the at least two parent electronic chromosomes using one or more randomly selected cut-points on the parent electronic chromosomes; and

means for evaluating the fitness value of the resulting child electronic chromosome using the fitness function to determine if elements in the child electronic chromosome correspond to the compound.